

IN THE CLAIMS:

This listing of the claims replaces all prior versions and listings of the claims.

Claim 1. (previously presented) Method for generating persistent annotations of multimedia content, comprising one or more repetitions of the following steps:

actively selecting examples of multimedia content to be annotated by a user, wherein the examples of multimedia content are selected based on at least one criterion for achieving a maximal disambiguation result such that only those examples which are most ambiguous are selected, wherein the at least one criterion includes a confidence level of the selected examples, the confidence level being inversely proportional to a distance of a new feature of the selected examples from a separating hyperplane in an induced higher dimensional feature space;

accepting input annotations from said user for said selected examples;

propagating said input annotations to other instances of multimedia content; and

storing said input annotations and said propagated annotations.

Claim 2. (original) The method of claim 1, wherein the step of actively selecting is performed using a selection technique selected from the group consisting of: deterministic and probabilistic.

Claims 3 and 4. (canceled)

Claim 5. (previously presented) The method of claim 1, wherein an optimization criterion for active selection includes information measures.

Claim 6. (previously presented) The method of claim 1, wherein the multimedia content comprises one or more types selected from the group consisting of: audio, video, Web pages, time series data, sensor data, and XML data.

Claim 7. (original) The method of claim 1, wherein the input annotations are created by a user with reference to a vocabulary.

Claim 8. (original) The method of claim 7, wherein the vocabulary contains one or more items selected from the group consisting of: terms, concepts, labels, and annotations.

Claim 9. (previously presented) The method of claim 1, wherein the process of creating input annotations by the user involves multimodal interaction with the user.

Claim 10. (original) The method of claim 1, wherein the input annotations are created by means of steps selected from the group consisting of: creating new annotations, deleting existing annotations, rejecting proposed annotations, and modifying annotations.

Claim 11. (original) The method of claim 7, wherein the vocabulary is adaptively or dynamically organized and/or limited by the system or the user.

Claim 12. (previously presented) The method of claim 9, wherein the multimodal interaction involves one or more elements selected from the group consisting of: gaze detection, finger pointing, expression detection, and effective computing methods for sensing a user's state.

Claim 13. (previously presented) The method of claim 1, wherein the propagation of annotations is determined deterministically or probabilistically and on the use of models for each annotation or for joint annotations.

Claim 14. (previously presented) The method of claim 13, wherein the models are created or learned automatically or semi-automatically and/or are updated adaptively from interaction with the user.

Claim 15. (previously presented) The method of claim 13, wherein the models are based on nearest neighbor voting or variants.

Claims 16-22. (canceled)

Claim 23. (previously presented) The method of claim 1, wherein the at least one criterion includes an ambiguity level of the selected examples.

Claims 24-28. (canceled)

Claim 29. (previously presented) The method of claim 13, wherein the models are based on expert systems.

Claim 30. (previously presented) The method of claim 13, wherein the models are based on rule-based systems.

Claim 31. (previously presented) The method of claim 13, wherein the models are based on hybrid techniques.

Claim 32. (previously presented) The method of claim 1, wherein the multimedia content comprises surveillance data.

Claim 33. (previously presented) Method for generating persistent annotations of multimedia content, comprising one or more repetitions of the following steps:

actively selecting examples of multimedia content to be annotated by a user, wherein the examples of multimedia content are selected based on at least one criterion for achieving a maximal disambiguation result such that only those examples which are most ambiguous are selected, wherein the at least one criterion is measured according to a model;

accepting input annotations from said user for said selected examples;

propagating said input annotations to other instances of multimedia content;

storing said input annotations and said propagated annotations; and

using said input annotations as training data to update the model.

Claim 34. (previously presented) The method of claim 33, further comprising repeating the step of selecting examples of multimedia content, wherein the repeated step of selecting examples of multimedia content is performed according to the model updated by using said input annotations as training data.

Claim 35. (previously presented) The method of claim 33, wherein the model uses at least one feature representation.

Claim 36. (previously presented) The method of claim 35, wherein the at least one feature representation comprises a texture.

Claim 37. (previously presented) The method of claim 35, wherein the at least one feature representation comprises a cepstral coefficient.

Claim 38. (previously presented) The method of claim 35, wherein the at least one feature representation comprises zero crossings.

Claim 39. (previously presented) The method of claim 33, further comprising performing user verification when the step of propagating said input annotations has been performed with least confidence.

Claim 40. (previously presented) The method of claim 33, wherein the at least one criterion includes a confidence level of the selected examples, the confidence level being inversely proportional to a distance of a new feature of the selected examples from a separating hyperplane in an induced higher dimensional feature space.

Claim 41. (previously presented) Method for generating persistent annotations of multimedia content, comprising one or more repetitions of the following steps:

actively selecting examples of multimedia content to be annotated by a user, wherein the examples of multimedia content are selected based on at least one criterion for achieving a maximal disambiguation result such that only those examples which are most ambiguous are selected;

accepting input annotations from said user for said selected examples, wherein a rectangular region of an image is associated with at least one of said input annotations;

propagating said input annotations to other instances of multimedia content; and

storing said input annotations and said propagated annotations.

Claim 42. (previously presented) The method of claim 41, wherein the at least one criterion includes a confidence level of the selected examples, the confidence level being inversely proportional to a distance of a new feature of the selected examples from a separating hyperplane in an induced higher dimensional feature space.